# How to generate a significant effective temperature for cold dark matter, from first principles

Patrick McDonald*

*Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, ON M5S 3H8, Canada*
(Dated: October 6, 2009)

I show how to reintroduce velocity dispersion into perturbation theory (PT) calculations of structure in the Universe, i.e., how to go beyond the pressureless fluid approximation, starting from first principles. This addresses a possible deficiency in uses of PT to compute clustering on the weakly non-linear scales that will be critical for probing dark energy. Specifically, I show how to derive a non-negligible value for the (initially tiny) velocity dispersion of dark matter particles, $\langle \delta v^2 \rangle$, where $\delta v$ is the deviation of particle velocities from the local bulk flow. The calculation is essentially a renormalization of the homogeneous (zero order) dispersion by fluctuations 1st order in the initial power spectrum. For power law power spectra with $n > -3$, the small-scale fluctuations diverge and significant dispersion can be generated from an arbitrarily small starting value – the dispersion level is set by an equilibrium between fluctuations generating more dispersion and dispersion suppressing fluctuations. For an $n = -1.4$ power law normalized to match the present non-linear scale, the dispersion would be $\sim 100$ km s$^{-1}$. This $n$ corresponds roughly to the slope on the non-linear scale in the real $\Lambda$CDM Universe, but $\Lambda$CDM contains much less initial small-scale power – not enough to bootstrap the small starting dispersion up to a significant value within linear theory (viewed very broadly, structure formation has actually taken place rather suddenly and recently, in spite of the usual "hierarchical" description). The next order PT calculation, which I carry out only at an order of magnitude level, should drive the dispersion up into balance with the growing structure, accounting for small dispersion effects seen recently in simulations.

## I. INTRODUCTION

Observing the large-scale density fluctuations in the Universe is one of the best ways we have to approach many fundamental questions about the Universe, e.g., understanding inflation, dark matter, dark energy, the curvature of the Universe, neutrino masses, possible extra dimensions, modifications of gravity, etc. [e.g., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16]. Statistics of the current density fluctuations can be used to infer statistics of the small initial perturbations from which they grew, and in turn to understand the physics of the very early Universe. Measuring the evolution of large-scale structure (LSS) over time tells us about the present matter content of the Universe and the dynamical rules its evolution follows. Before we can learn anything about fundamental properties of the Universe, however, we must be able to compute the directly observable astrophysical quantities (e.g., the CMB [17], galaxy clustering statistics [18], Ly$\alpha$ forest absorption [19, 20, 21, 22, 23, 24], galaxy ellipticity correlations used to measure weak lensing [25], galaxy cluster/Sunyaev-Zel'dovich effect (SZ) measurements [26], and possibly future 21cm surveys [27]) given a hypothetical underlying model.

As observational statistics become more and more precise, with the potential to measure more and more subtle differences between models, the requirements on the phenomenological theory needed for their interpretation become correspondingly more stringent. Currently, linear-order perturbation theory [28, 29, 30, 31] provides our primary means of calculating LSS observables for cosmological parameter estimation and model testing, with only ad hoc, and recently demonstrably inadequate, attempts to correct for non-linearity once it becomes non-negligible [1, 8, 18, 32, 33] (a vast number of papers have been written on beyond-linear calculations, but most of these are never used in parameter estimation papers). Linear theory can robustly describe observations on very large scales or at very early times, but breaks down when the perturbations become too large on a given scale. When considering gravitational evolution only (i.e., dark matter only), numerical simulations can be used to compute the fully non-linear evolution of the density field to high accuracy (with a lot of care and computer power [34, 35, 36]), however, as discussed extensively in [37], numerical simulations are a less than ideal tool for interpreting future precision observations, once one considers real observables which are inevitably influenced by baryonic effects (e.g., star formation, and the accompanying complication of general gas dynamics). To summarize the argument in [37]: Beyond linear order perturbation theory (PT) [29, 38, 39, 40, 41, 42, 43, 44] should provide the primary means of interpreting very high precision future

---

LSS data, just like linear theory has provided the primary means in the past. The range of scales over which higher order PT will be necessary (i.e., linear theory is insufficient) and sufficient (i.e., it will be accurate enough after computing a modest number of terms) will become larger as the precision of observations increases, while the chances of robustly, completely, convincingly describing the full precision of the observations using inevitably somewhat ad hoc prescriptions for star formation and related things in simulations becomes more remote. The key difference between PT and simulations is the fact that perturbative clustering can be completely described by a finite set of well-defined parameters, no matter how complicated the small-scale physics is (at least as conjectured in [37]), while the need for fully non-linear calculations implies that there is generally no bound on the number or form of free parameters (more or less by definition). The idea that the importance of PT relative to simulations is increasing with time may seem backwards relative to conventional wisdom, however, my argument is that this conventional wisdom was developed for the era of not very precise observations, when corrections to linear theory were already too large to be described by PT by the time they were large enough to be statistically significant, i.e., past PT work was premature. To be clear, I am not saying that simulations will not be extremely useful, only that they are unlikely to be the leading tool for extracting fundamental cosmological information from very high precision observations (much like the situation in high energy physics, where lattice QCD simulations provide much qualitative, and recently even quite precise quantitative, insight [45], but high precision constraints on models are made primarily in the regime accessible to perturbation theory [46]).

Following the above line of reasoning, this paper is aimed at building up our ability to do calculations based on perturbation theory. It deals exclusively with the gravity-only part of the calculation, however, this should be viewed only as an intermediate goal. PT is most necessary for practical computations of observables which can not realistically be done from first principles in a simulation – understanding how to do computations for dark matter-only is simply a prerequisite for computing these observables. This continues a recent line of work related to renormalization/resummation methods [47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63]. In my opinion, the key to maximizing the usefulness of all of this work is eventually connecting it to galaxy biasing and related complications [37, 64, 65, 66].

One traditional potential limitation of LSS perturbation theory, which I focus on addressing in this paper, is that it assumes the particles at a point in space all have exactly the same velocity [67]. The equations used are those of a pressureless fluid. This is often referred to as the "single-stream approximation", however, I will avoid this language as it assumes a certain picture of large-scale structure formation that may or may not have anything to do with reality. I say this because, for typically observed times, stream crossing is actually ubiquitous – in fact, there are many orders of magnitude in scale of deeply non-linear structure, to the point where the initial conditions on very small scales may be effectively forgotten. Of course, the standard language implicitly means no stream crossing when the field is in some sense smoothed on the typical scale where the perturbation theory is supposed to apply. When you look at it this way, however, it is clear that, if coldness is a good effective theory, it is not simply because the particles were initially cold, there is an additional requirement that the velocity field remains effectively cold on scales just smaller than the ones of interest.

To give a qualitative preview of the paper: The exact equation for the evolution of collisionless particles is the Vlasov equation for the full distribution function in 6-dimensional phase space (particle density in position and momentum space). The standard hydrodynamic equations are derived by taking moments of the Vlasov equation with respect to momentum – the zeroth moment gives an equation for the evolution of density, the first moment gives an equation for the evolution of bulk velocity, and higher moments are normally dropped, e.g., the second moment which would describe velocity dispersion. At first glance, one might think that standard PT could be extended by simply adding the evolution equation for the 2nd and possibly higher moments, however, that turns out to be less straightforward than it sounds. Viewed conventionally, dropping these higher moments is not really an added approximation in standard Eulerian perturbation theory. The lowest order evolution equations contain only a Hubble drag term, so any small initial velocity dispersion will rapidly become even smaller. Higher order equations contain no source terms, meaning that the higher order results are always proportional to the tiny starting value. [68] showed that even perturbation theory using the full distribution function directly leads to the same result. I will show, however, that this argument for dropping dispersion from the perturbation theory is faulty, in that, while the lowest order terms are very small, the series is very rapidly diverging, in the sense that higher order terms are increasing in size instead of decreasing. This implies that some rearrangement of the series is needed, as in a resummation or renormalization group calculation. In case this mathematical reasoning is not sufficient motivation, recently [67] computed the velocity dispersion generated in N-body simulations, finding it to be small but not completely dynamically negligible.

There has been much discussion in the past about different ways of adding velocity dispersion, or more general changes to the small-scale effective theory used for PT calculations [e.g., 69, 70, 71, 72, 73, 74, 75], however, these papers all lacked a first principles derivation, starting from the exact equations, of the model they use. This made their usefulness for precision calculations questionable, as there were always added assumptions and/or free parameters. The point of this paper is to show how to do a straightforward computation that takes us directly from the initial

homogeneous, cold, starting theory to a theory with highly developed, potentially hot, small-scale structure.

The rest of the paper is as follows: In §II I show how to use renormalization group-inspired ideas to reintroduce velocity dispersion from first principles. In §III I give a short discussion of the implications of these velocity dispersion calculations for redshift-space distortions. Finally, in §IV I discuss the results.

## II. GENERATION OF VELOCITY DISPERSION

In this section I present the calculation that generates velocity dispersion. In §II A I discuss the time evolution equations for cold dark matter, which are the starting point for perturbation theory. In §II B I compute the velocity dispersion taking the standard PT approach, which leads to negligible dispersion. In §II C I apply an approximate renormalization group approach to models with power law initial power spectra, which demonstrates how significant dispersion can be generated. Finally, in §II D I take a related but more exact approach, which should be easily extendible to more general, higher order, calculations, including ΛCDM power spectra.

### A. Evolution equations

The exact evolution of collisionless particles is is described by the Vlasov equation [29]

$$\frac{\partial f}{\partial \tau} + \frac{1}{a\,m}\,\mathbf{p} \cdot \nabla f - a\,m\,\nabla \phi \cdot \nabla_{\mathbf{p}} f = 0 \ , \tag{1}$$

with

$$\nabla^2 \phi = 4\,\pi\,G\,a^2 \bar{\rho}\,\delta = \frac{3}{2} H_0^2\,\Omega_{m,0}\,\frac{\delta}{a} \ , \tag{2}$$

where $f(\mathbf{x}, \mathbf{p}, \tau)$ is the particle density at phase-space position $(\mathbf{x}, \mathbf{p})$, $m$ is the particle mass (which plays no role in the final results), and $\mathbf{p} = a\,m\mathbf{v}_p$ ($\mathbf{v}_p$ here is a particle's peculiar velocity, not to be confused with the mean peculiar velocity used everywhere else). $\mathbf{x}$ is the comoving position and $\tau = \int dt/a$ is the conformal time, with $a = 1/(1+z)$ the expansion factor. Except when otherwise indicated $\nabla = \partial/\partial \mathbf{x}$. The density field is obtained by averaging the distribution function over momentum:

$$\rho(\mathbf{x}, \tau) \equiv ma^{-3} \int d^3 p\ f(\mathbf{x}, \mathbf{p}, \tau) \ , \tag{3}$$

and the bulk (mean) velocity and higher moments of the velocity distribution e.g., the dispersion of particle velocities around their bulk velocity, can be similarly obtained by multiplying the distribution function by any number of $\mathbf{p}$'s (e.g., one to obtain bulk velocity) before integrating over $\mathbf{p}$. The mean velocity of the particles at $\mathbf{x}$ is

$$\mathbf{v}\,(\mathbf{x}, \tau) \equiv \frac{\int (\mathbf{p}/ma)\,f\ d^3 p}{\int f\ d^3 p} \ . \tag{4}$$

and the velocity dispersion tensor is

$$\sigma^{ij}\,(\mathbf{x}, \tau) \equiv \frac{\int \left(p_i p_j / m^2 a^2\right) f\ d^3 p}{\int f\ d^3 p} - v^i v^j \ , \tag{5}$$

i.e., $\sigma^{ij}(\mathbf{x}, \tau) \equiv \left\langle \delta v^i\ \delta v^j \right\rangle_p$ with $\delta v^i$ the deviation of a particle's velocity from the local mean velocity, and the average is over all particles at $\mathbf{x}$.

As discussed by [29], taking moments of the Vlasov equation with respect to momentum leads to a hierarchy of evolution equations for these quantities. The density evolution equation is the usual continuity equation,

$$\frac{\partial \delta}{\partial \tau} + \partial_i \left[(1 + \delta)\,v^i\right] = 0 \ , \tag{6}$$

where $\delta(\mathbf{x}, \tau) = \rho(\mathbf{x}, \tau)/\bar{\rho} - 1$, $\partial_k = \partial/\partial x^k$. The bulk velocity evolution equation acquires a velocity dispersion term that is usually dropped to give the Euler equation for standard perturbation theory

$$\frac{\partial v^i}{\partial \tau} + \mathcal{H}v^i + v^j \partial_j v^i = -\partial_i \phi - \frac{\partial_j \left[(1 + \delta)\,\sigma^{ij}\right]}{1 + \delta} \ , \tag{7}$$

where $\mathcal{H} = d\ln a/d\tau = aH$ with $H$ the usual Hubble parameter. Multiplying the Vlasov equation by $p_i p_j$ and integrating over $\mathbf{p}$ gives (after some substitution to remove bulk velocity terms)

$$\frac{\partial \sigma^{ij}}{\partial \tau} + 2\,\mathcal{H}\,\sigma^{ij} + v^k\,\partial_k \sigma^{ij} + \sigma^{jk}\,\partial_k v^i + \sigma^{ik}\,\partial_k v^j = -\frac{\partial_k \left[(1+\delta)\,q^{ijk}\right]}{1+\delta}\,, \tag{8}$$

where $q^{ijk}(\mathbf{x},\tau) \equiv \left\langle \delta v^i\,\delta v^j\,\delta v^k \right\rangle_p$. Similar equations can be derived for $q^{ijk}$ and higher moments.

## B. Standard perturbation theory

Perturbation theory consists of writing the fields as a series of terms of at least formally increasing order of smallness, i.e., $\delta = \delta_1 + \delta_2 + \delta_3 + \dots$. The evolution equations are solved order-by-order, with lower order solutions appearing as sources in the higher order equations so that $\delta_n$ is of order $\delta_1^n$ [43].

For simplicity, I will often assume an Einstein-de Sitter (EdS) Universe. This assumption makes analytic calculations easier without qualitatively changing the results. We then have useful relations:

$$a = \left(\frac{\tau}{3t_0}\right)^2 = \left(\frac{\tau H_0}{2}\right)^2,\quad \mathcal{H} = \frac{2}{\tau},\quad \text{and}\quad \frac{3}{2}\frac{H_0^2}{a} = \frac{6}{\tau^2}\,. \tag{9}$$

Symmetry allows for a zero order (assumed to be homogeneous and isotropic) component of $\sigma^{ij}$, $\bar{T}_0(\tau) \equiv \bar{\sigma}_0^{11} = \bar{\sigma}_0^{22} = \bar{\sigma}_0^{33}$, and in fact we expect there to be primordial velocity dispersion for realistic WIMPs, albeit very small. The background dispersion evolves following:

$$\bar{T}_0(\tau) \propto a^{-2} \propto \tau^{-4}\,, \tag{10}$$

where the normalization must be fixed by the initial conditions.

The linearized equations are:

$$\frac{\partial \delta_1}{\partial \tau} + \partial_i v_1^i = 0 \tag{11}$$

$$\frac{\partial v_1^i}{\partial \tau} + \mathcal{H} v_1^i = -\partial_i \phi_1 - \partial_j \sigma_1^{ij} - \bar{T}_0 \partial_i \delta_1 \tag{12}$$

$$\frac{\partial \sigma_1^{ij}}{\partial \tau} + 2\,\mathcal{H}\,\sigma_1^{ij} + \bar{T}_0\,\partial_j v_1^i + \bar{T}_0\,\partial_i v_1^j = 0\,, \tag{13}$$

where the Poisson equation holds order by order relating $\delta$ to $\phi$. Here is where the story ends for velocity dispersion in standard PT. The CDM velocity dispersion is supposed to be initially small, and only gets smaller. At 1st order the evolution equation for $\sigma^{ij}$ contains only the Hubble drag term and terms proportional to the tiny $\bar{T}_0$, so the perturbations will remain small. The last term from Eq. (7) can be dropped and the hierarchy is closed. I will show that this treatment is justified in the approach of standard PT, but retaining the apparently small dispersion terms allows for the renormalization in the next subsection. I will drop $q^{ijk}$ for simplicity (and because symmetry prevents it from acquiring a non-zero background value).

In this paper my only goal is to renormalize the zero-order velocity dispersion, $\bar{T}_0$, so I am going to henceforth assume the velocity field is curl-free and only solve for $\nabla \cdot \mathbf{v}$ and $\partial_i \partial_j \sigma^{ij}$. The techniques of this paper could probably be used to reactivate the vorticity variables, but symmetry guarantees that they will not have any homogeneous (zero order, i.e., background) component (see [67] for a measurement of the vorticity power spectrum in simulations, and an argument that vorticity is completely irrelevant for large-scale clustering). The linearized equations, assuming EdS, can be re-written as:

$$\frac{\partial \delta_1}{\partial \eta} = \theta_1 \tag{14}$$

$$\frac{\partial \theta_1}{\partial \eta} + \frac{1}{2}\,\theta_1 = \frac{3}{2}\,\delta_1 - 2\,k^2 \tilde{T}_0\,(\delta_1 + 2\pi_1) \tag{15}$$

$$\frac{\partial \pi_1}{\partial \eta} + \pi_1 = \theta_1 - \frac{\partial \ln \tilde{T}_0}{\partial \eta}\,\pi_1\,, \tag{16}$$

where

$$\eta = \ln a \tag{17}$$

$$\theta = \frac{i\mathbf{k}\cdot\mathbf{v}}{\mathcal{H}} \tag{18}$$

$$\pi = \frac{1}{2}\frac{k_i k_j}{k^2}\frac{\sigma^{ij}}{\bar{T}_0} \tag{19}$$

$$\tilde{T}_0 = \frac{\bar{T}_0}{\mathcal{H}^2} \tag{20}$$

I have moved to Fourier space, e.g., $\delta(\mathbf{k},\tau) = \int d^3\mathbf{x}\ \exp(i\mathbf{k}\cdot\mathbf{x})\ \delta(\mathbf{x},\tau)$, where $\nabla \rightarrow -i\mathbf{k}$.

These equations can't be solved exactly. We expect that the density and velocity modes will be as usual linear theory at low $k$ until suppressed below some Jeans-like filtering scale. The basic behavior can be found more quantitatively from a small $k$ expansion, where I first drop the $k^2$ terms and obtain, using $\tilde{T}_0 = \tilde{T}_i a_i/a$, the usual growing mode solutions, $\delta_1(k\rightarrow 0) = \theta_1(k\rightarrow 0) = \pi_1(k\rightarrow 0) = \delta_i a/a_i$. I then substitute these solutions into the $k^2$ terms to obtain, to order $k^2$,

$$\delta_1 = \theta_1 = \pi_1 = \delta_i\frac{a}{a_i}\left(1 - \frac{6}{5}k^2\tilde{T}_i\right)\ . \tag{21}$$

Here I have eliminated the new integration constants at order $k^2$ by requiring that the $\mathcal{O}(k^2)$ corrections are zero at some very early initial time $a_i$. With the assumption that the fields are curl free, we can invert the $\pi_1$ solution to obtain

$$\sigma_1^{ij} = 2\bar{T}_0\frac{k_i k_j}{k^2}\delta_1 \tag{22}$$

If I assume the Jeans-like suppression kernel is Gaussian, i.e., $\exp\left(-k^2 R_F^2/2\right)$, we have $1 - 6k^2\tilde{T}_i/5 = 1 - k^2 R_F^2/2$ so $R_F = (12\tilde{T}_i/5)^{1/2}$. Note that I have in this derivation only retained the fasted growing parts of $\delta_1$, $\theta_1$, and $\pi_1$, which is consistent in standard perturbation theory where the initial time can be taken to be arbitrarily small, but, as we will see, is dangerous when we start renormalizing. Finally, note that as long as $k^2\tilde{T}_i$ is very small, which it will be for CDM on observable scales (basically by definition), we have changed nothing by retaining the velocity dispersion terms, i.e., the standard perturbation theory with zero velocity dispersion is self-consistent at this order.

The 2nd order equations are:

$$\frac{\partial\delta_2}{\partial\tau} + \partial_i v_2^i + \delta_1\partial_i v_1^i + v_1^i\partial_i\delta_1 = 0 \tag{23}$$

$$\frac{\partial v_2^i}{\partial\tau} + \mathcal{H}v_2^i + v_1^j\partial_j v_1^i = -\partial_i\phi_2 - \partial_j\sigma_2^{ij} + \delta_1\bar{T}_0\partial_i\delta_1 - \sigma_1^{ij}\partial_j\delta_1 - \bar{T}_0\partial_i\delta_2 \tag{24}$$

$$\frac{\partial\sigma_2^{ij}}{\partial\tau} + 2\ \mathcal{H}\ \sigma_2^{ij} + \bar{T}_0\ \partial_j v_2^i + \bar{T}_0\ \partial_i v_2^j = -v_1^k\ \partial_k\sigma_1^{ij} - \sigma_1^{jk}\ \partial_k v_1^i - \sigma_1^{ik}\ \partial_k v_1^j\ . \tag{25}$$

For now I am only interested in $\bar{T}_2 \equiv \langle\sigma_2^{11}\rangle = \langle\sigma_2^{22}\rangle = \langle\sigma_2^{33}\rangle = \langle\sigma_2^{ii}\rangle/3$, which will renormalize $\bar{T}_0$.

$$\frac{\partial\bar{T}_2}{\partial\tau} + 2\ \mathcal{H}\ \bar{T}_2 + \frac{1}{3}\left\langle v_1^k\ \partial_k\sigma_1^{ii} + 2\sigma_1^{ik}\ \partial_k v_1^i\right\rangle = 0\ . \tag{26}$$

Evaluating this in Fourier space, using the same variable redefinitions as above, including $\tilde{T}_2 = \bar{T}_2\mathcal{H}^{-2}$, gives

$$\frac{\partial\tilde{T}_2}{\partial\eta} + \tilde{T}_2 = \frac{2}{3}\tilde{T}_0\langle\pi_1\theta_1\rangle \simeq \frac{2}{3}\tilde{T}_0\langle\delta_1^2\rangle\ , \tag{27}$$

where deriving the term involving $\langle\pi_1\theta_1\rangle$ only requires that $\theta_1$ and $\pi_1$ completely describe the velocities, while the term involving $\langle\delta_1^2\rangle$ requires the above approximate relation between $\delta_1$, $\theta_1$, and $\pi_1$.

Using the right-most side of Eq. (27), and assuming standard linear theory for $\delta$, the solution is:

$$\tilde{T}_2 = \frac{1}{3}\tilde{T}_0\langle\delta_1^2\rangle + \frac{c}{a} \tag{28}$$

where $c$ is a constant. If I choose $c$ to make $\tilde{T}_2(a = a_i) = 0$, and use the zero order solution $\tilde{T}_0 = A/a$ (remembering that $\tilde{T} = \bar{T}/\mathcal{H}^2$), I finally obtain

$$\tilde{T}(a) = \frac{A}{a} + \frac{1}{3}\left[\frac{A}{a}\left\langle\delta_1^2\right\rangle - \frac{A}{a_i}\left\langle\delta_1^2\right\rangle_i\frac{a_i}{a}\right]. \tag{29}$$

The standard PT approach would be to drop the term on the right in the brackets, assuming $a >> a_i$. From the point of view of standard PT, we could conclude that the result for $\tilde{T}_2$ is small, for $\Lambda$CDM, so dropping velocity dispersion is justified (while $\left\langle\delta_1^2\right\rangle$ is large for $\Lambda$CDM, it is not large enough to overcome the small dispersion and make $\left\langle\delta_1^2\right\rangle\tilde{T}$ large); however, we have a clear breakdown in the premise of the perturbation theory, because $\tilde{T}_2/\tilde{T}_0 \sim \left\langle\delta_1^2\right\rangle >> 1$.

## C.    Renormalization group approach

The problem with the standard calculation outlined in §II B, which leads me to renormalization, is that the 2nd term in Eq. (29) ($\tilde{T}_2$) diverges relative to the first ($\tilde{T}_0$), increasingly as time progresses. One might argue that the correction is still small, in an absolute sense, but this is nonsense because there is no reason not to expect the higher order terms in the series to be even larger, i.e., the truth could be arbitrarily large. Fortunately, we have tools to deal with this kind of breakdown in perturbative solutions of differential equations [76, 77, 78, 79, 80].

A key observation about perturbation theory is that there is inevitable ambiguity in the solution that comes from solving a set of differential equations at each order. We obtain a new set of integration constants at each order, while the initial conditions only determine one set. This is apparent in Eq. (29), where I fixed the 2nd order integration constant by the *arbitrary* requirement that $\tilde{T}_2(a = a_i) = 0$. I can instead fix it to set $\tilde{T}_2(a_\star) = 0$, where $a_\star$ is some arbitrary time, so the solution is

$$\tilde{T}(a) = \frac{A_\star}{a} + \frac{1}{3}\frac{A_\star}{a}\left[\left\langle\delta_1^2\right\rangle - \left\langle\delta_1^2\left(a_\star\right)\right\rangle\right] \tag{30}$$

where $A_\star$ must depend on $a_\star$, in order to match the initial conditions. The final result should not depend on the arbitrary $a_\star$, so the RG equation is obtained by taking the derivative of $\tilde{T}(a)$ with respect to $a_\star$, evaluated at $a = a_\star$, and setting it equal to zero:

$$\left.\frac{d\tilde{T}}{da_\star}\right|_{a_\star=a} = 0 = \frac{dA_\star}{da}a^{-1} - \frac{1}{3}A_\star a^{-1}\frac{\partial\left\langle\delta_1^2\right\rangle}{\partial a} \tag{31}$$

or

$$\frac{dA_\star}{da} = \frac{1}{3}A_\star\frac{\partial\left\langle\delta_1^2\right\rangle}{\partial a} . \tag{32}$$

Now, $\left\langle\delta_1^2\right\rangle$ generally depends on $A_\star$ through the filtering scale $R_F$, as discussed above, but supposed for the moment that the power spectrum was not truncated so $\left\langle\delta_1^2\right\rangle$ is independent of $A_\star$. The solution to the RG equation would be, long after the initial time,

$$\frac{A_\star}{A_i} = \exp\left(\frac{\left\langle\delta_1^2\right\rangle}{3}\right) \tag{33}$$

where $A_i$ is the initial $A$. For a power law with $n > -3$, $\left\langle\delta_1^2\right\rangle$ is infinite, so we have infinite growth of the velocity dispersion, but even in the case of $\Lambda$CDM with asymptotic high-$k$ slope slightly less than -3 (because the primordial slope, from inflation, appears to be less than one [1]), $\exp\left(\left\langle\delta_1^2\right\rangle/3\right)$ will become enormous. Furthermore, as discussed in [55], higher order corrections to the power spectrum generically push the slope to larger values, so the result, without filtering, will be even larger when calculated to higher order, to the point of being practically infinite. Infinite velocity dispersion is of course too much – the point here is simply to demonstrate how the ridiculously small initial seed velocity dispersion in CDM can grow into substantial velocity dispersion later.

Including the Jeans filtering resulting from the velocity dispersion itself will regulate the growth of the dispersion. Recall that the approximate filtering scale is $R_F = (12\tilde{T}_i/5)^{1/2} = (12A/5a_i)^{1/2}$. To be precise, $R_F$ quantifies the frozen-in power suppression due to some velocity dispersion present at initial time $a_i$, long after this dispersion has redshifted away. When we in effect add new dispersion at time $a_\star$, through the RG equation, the smoothing will

not be instantaneous; however, it takes place quite quickly, so I will assume I can use $R_F = [12A_\star/5\alpha a_\star]^{1/2}$ in the RG equation, where $\alpha \sim 3$ is a fudge factor to account for the lag between the introduction of dispersion and the smoothing of the power spectrum. As with the original definition for $R_F$, where I assumed that the smoothing is Gaussian, this approximation means the results should only be trusted at the order-of-magnitude level. I will do a more exact calculation below. Assuming Gaussian filtering, and power spectrum $\Delta^2(k, a) = \Delta_p^2(a)(k/k_p)^{3+n}$ (with $n > -3$), the variance is

$$\langle \delta_1^2 \rangle = \frac{\Delta_p^2(a)}{2} \frac{\Gamma\left[(3+n)/2\right]}{\left(k_p R_F\right)^{3+n}} \ . \tag{34}$$

Using Eq. (34) in Eq. (32) I find, well after the initial time,

$$A = \left[\frac{3+n}{6} \frac{\Delta_p^2(a)}{2} \frac{\Gamma\left((3+n)/2\right)}{\left(k_p \sqrt{12/5\alpha a}\right)^{3+n}}\right]^{\frac{2}{3+n}} . \tag{35}$$

Recall that $\tilde{T}(a) = Aa^{-1}$. This result may not be very illuminating, but it can be re-written in a way that makes it very clear:

$$\frac{6}{3+n} = \left\langle \delta_1^2\left[R_F(A)\right]\right\rangle \ , \tag{36}$$

i.e., the velocity dispersion simply grows to the point where the Jeans-like filtering reduces the variance to of order unity, with the exact relation dependent on the slope of the power spectrum.

### 1. RG method, 2nd iteration

The calculation leading to Eq. (36) does not look entirely self-consistent. At various stages in it, I used the fact that $\tilde{T}_0 \propto a^{-1}$, however, in the end, the renormalized $\tilde{T}$ is $\propto a^{4/(3+n)}$, growing quickly rather than decaying. The brute force way to solve this problem would have been to compute RG equations for the pieces of the calculation that depended on this assumption (e.g., the amplitude of $\pi$), and solve them jointly. It is simpler, and will lead to self-consistent and enlightening results, to redo the calculation starting with $\tilde{T}_0 = Aa^\alpha$, with $\alpha = 4/(3+n)$. The large-scale solutions for $\delta$ and $\theta$ are unchanged, but we now have $\pi_1(k \to 0) = (2+\alpha)^{-1}\delta_1(k \to 0)$. The smoothing estimated from the $k^2$ expansion changes for all three fields:

$$\delta_1 \simeq \left[1 - \frac{2}{(2+\alpha)}k^2\tilde{T}_0\frac{(4+\alpha)}{\alpha(5+2\alpha)}\right]\delta(k \to 0) \tag{37}$$

$$\theta_1 \simeq \left[1 - \frac{2(1+\alpha)}{(2+\alpha)}k^2\tilde{T}_0\frac{(4+\alpha)}{\alpha(5+2\alpha)}\right]\delta(k \to 0) \tag{38}$$

$$\pi_1 \simeq \frac{1}{2+\alpha}\left[1 - k^2\tilde{T}_0\frac{(4+\alpha)}{\alpha(5+2\alpha)}\right]\delta(k \to 0) \ . \tag{39}$$

Note that there is a qualitative difference in these smoothing kernels relative to the $\alpha < 0$ case in that $\tilde{T}$ appears instead of $\tilde{T}_i$, i.e., because the comoving Jeans scale is increasing when $\alpha > 0$, the smoothing continues to grow with time rather than freezing out. The numerical factors are also different, with $\theta_1$ being generally significantly smoother than $\delta_1$, with $\pi_1$ between them. For example, for $n = -1.4$, $R_F^\delta \simeq 0.48\,\tilde{T}_0^{1/2}$, $R_F^\theta \simeq 0.90\,\tilde{T}_0^{1/2}$, and $R_F^\pi \simeq 0.72\,\tilde{T}_0^{1/2}$. Now,

$$\langle \pi_1\theta_1 \rangle = \frac{1}{2+\alpha}\frac{\Delta_p^2(a)}{2}\frac{\Gamma\left[(3+n)/2\right]}{\left(k_p R_F^{\pi\theta}\right)^{3+n}} \tag{40}$$

where $R_F^{\pi\theta 2} = (R_F^{\pi 2} + R_F^{\theta 2})/2$ and $R_F^\theta$ and $R_F^\pi$ are the smoothing scales implied by Eq. (38) and (39). The important thing to recognize about $\sigma_{\pi\theta}^2 = \langle \pi_1\theta_1 \rangle$ here is that it is time independent, i.e., the growth of the power spectrum is canceled by the change in $R_F^{\pi\theta 2} \propto \tilde{T}_0$.

Now, I have recomputed the linear equations given $\tilde{T}_0 \propto a^\alpha$, so it remains only to recompute the perturbative corrections $\tilde{T}_2$. Eq. (27) for $\tilde{T}_2$ is changed, because $\tilde{T}_0$ is no longer a solution to the zero order equation for $\tilde{T}$. This is a key fact – at the heart of all renormalization is the concept that the naive lowest order result is not always the best one to perturb around when doing computations to higher order. Sometimes it is better to perturb around something else, with the criteria for "better" being simply that the corrections remain small. The usual renormalization procedure of doing the calculation to lowest order, using the results to compute higher order corrections, then absorbing large corrections back into the lowest order, is an algorithmic way of accomplishing what one could accomplish by simply perturbing around the better starting point from the beginning (if one has a way to determine/guess what it is). Here, I am perturbing around $\tilde{T}_0 = A a^\alpha$, i.e., $\tilde{T}_2$ is defined by $\tilde{T} = A a^\alpha + \tilde{T}_2 + ...$, and the evolution equation I derive for $\tilde{T}_2$ is:

$$\frac{\partial \tilde{T}_2}{\partial \eta} + \tilde{T}_2 = \frac{2}{3} \tilde{T}_0 \langle \pi_1 \theta_1 \rangle - (\alpha + 1) \tilde{T}_0 \ . \tag{41}$$

The 2nd term on the right hand side is new. Note that I have not yet fixed $A$. Normally, $A$ would be set by the initial conditions, but there is no need for that because the homogeneous part of the solution for $\tilde{T}_2$ can be used to match the initial conditions (then, since the homogeneous solution is just $\propto a^{-1}$, the memory of initial conditions will fade away). If I choose $A$ to make $\frac{2}{3} \langle \pi_1 \theta_1 \rangle - (\alpha + 1) = 0$, which is possible because $\sigma_{\pi\theta}^2$ depends on $A$ through the smoothing kernels, the equation for $\tilde{T}_2$ becomes trivial, just the homogeneous equation, i.e., there are no perturbative corrections to $\tilde{T}_0 = A a^\alpha$! (That is, for the correct amplitude, and the special value $\alpha = 4/(3+n)$.) This result should not be viewed as some kind of fortuitous coincidence, or artificial tuning. It is what we aim for in a renormalization group calculation, and reflects the physical correctness of the idea that the velocity dispersion should be rapidly growing, tracking the non-linear scale, with the the unstable cold start completely irrelevant (at least after enough time has passed, and for this kind of power law power spectrum).

We immediately have an equation like Eq. (36), except coming from a much more accurate calculation:

$$\langle \pi_1 \theta_1 \rangle = \frac{3}{2} (\alpha + 1) = \frac{6}{3 + n} + \frac{3}{2} \ . \tag{42}$$

The variance of $\delta_1$ is larger, because of the small coefficient of $\pi_1$, and lesser smoothing of $\delta_1$, i.e., the more accurate version of Eq. (36) is:

$$\langle \delta_1^2 \left[ R_F^\delta (A) \right] \rangle = (2 + \alpha) \left( \frac{R_F^{\pi\theta}}{R_F^\delta} \right)^{3+n} \langle \pi_1 \theta_1 \rangle = 3 \frac{(5+n)(7+n)}{(3+n)^2} \left( \frac{6+n}{3+n} \right)^{(3+n)/2} \ , \tag{43}$$

e.g., this evaluates to $\sigma_\delta^2 = 55$ at $n = -1.4$, with $\sigma_{\pi\theta}^2 = 5.2$. Keep in mind that this is still essentially the linear theory variance. There will be non-linear corrections, however, especially for relatively large $n$, it seems like a very good thing to be starting with finite linear variance rather than the infinite variance of the bare power law. For example, for $n > -1$ the standard 1-loop PT correction diverges [42], while here that problem is clearly solved, i.e., we have a natural high-$k$ cutoff. Choosing the pivot point at $k_{\rm NL}$, defined by $\Delta^2(k_{\rm NL}) = 1$, these equations can be solved to give

$$k_{\rm NL} R_F^\delta = \left( \frac{\Gamma\left[ (3+n)/2 \right]}{2 \ \sigma_\delta^2} \right)^{\frac{1}{3+n}} = \left( \frac{(3+n)^2 \ \Gamma\left[ (3+n)/2 \right]}{6 (5+n)(7+n)} \right)^{\frac{1}{3+n}} \left( \frac{3+n}{6+n} \right)^{1/2} \tag{44}$$

or

$$k_{\rm NL} \tilde{T}_0^{1/2} = \left( \frac{(3+n)^2 \ \Gamma\left[ (3+n)/2 \right]}{6 (5+n)(7+n)} \right)^{\frac{1}{3+n}} \left( \frac{(5+n)(23+5n)}{2 (3+n)(4+n)(6+n)} \right)^{1/2} \ , \tag{45}$$

e.g., $k_{\rm NL} \tilde{T}_0^{1/2} = 0.12$ for $n = -1.4$. As one might guess, the velocity dispersion scale increases relative to the non-linear scale as the power law increases to include more small-scale power. Note that this calculation was self-consistent, without any very ugly approximations, although the use of Gaussian smoothing kernels based on the $k^2$ expansion makes the results still only approximate.

### D.   A more exact, general approach

A simpler way approach the split between homogeneous background and perturbations, which I could have just started with (except I think the above derivation has some pedagogical value), is to define $\sigma_{ij} (\tau, \mathbf{x}) =$

$\left\langle \delta v^i \ \delta v^j \right\rangle_p (\tau, \mathbf{x}) - \bar{T}(\tau)\, \delta^K_{ij}$, with $\bar{T}(\tau)\, \delta^K_{ij} \equiv \left\langle \left\langle \delta v^i \ \delta v^j \right\rangle_p (\tau, \mathbf{x}) \right\rangle_{\mathbf{x}}$. Then it is clear that

$$\frac{\partial \bar{T}}{\partial \tau} + 2\,\mathcal{H}\,\bar{T} = -\frac{1}{3} \left\langle v^k\ \partial_k \sigma^{ii} + 2\sigma^{ik}\ \partial_k v^i \right\rangle \ , \tag{46}$$

and

$$\frac{\partial \sigma^{ij}}{\partial \tau} + 2\,\mathcal{H}\,\sigma^{ij} + \bar{T}\left(\partial_i v^j + \partial_j v^i\right) = -\Delta\left[v^k\ \partial_k \sigma^{ij} + \sigma^{jk}\ \partial_k v^i + \sigma^{ik}\ \partial_k v^j\right] \tag{47}$$

where $\Delta[f] \equiv f - \langle f \rangle$ and I have dropped the $q^{ijk}$ term. These are the fully non-linear equations, with the need to include the right hand side of Eq. (46) as the source of homogeneous velocity dispersion, while subtracting the same thing from the perturbations, being a simple result of the definition of the mean and perturbations, rather than some kind of renormalization. Either way, however, we have the same basic idea that the perturbations feed back on the background, new relative to standard PT with just density and velocity.

### 1. RG approach at the level of equations rather than solutions

We still need some perturbative approach to solving the coupled equations for $\bar{T}$ and the fluctuations in density, velocity, and velocity dispersion. The RG approach used above and in [55] is not ideal. First, one needs to solve the perturbative equations analytically, which isn't generally possible without approximations that can lead to uncontrolled errors. Second, the RG equation that is derived may not be analytically solvable, as I found in [55]. We can solve the first problem (needing analytic solutions for the perturbation evolution), without necessarily making the second problem any worse, by considering applying the ideas behind the RG method at the level of equations rather than solutions. Suppose we want to solve

$$\dot{\delta} = f(\delta) \tag{48}$$

The standard perturbative equations are

$$\dot{\delta}_1 = f' \delta_1 \ , \qquad \dot{\delta}_2 = f' \delta_2 + \frac{f''}{2} \delta_1^2 \ , \qquad ... \tag{49}$$

etc., where $f' = df/d\delta(\delta = 0)$, etc.. Our solution up to 2nd order is $\delta_1 + \delta_2$, but note that we could solve a different pair of equations like this:

$$\dot{\delta}_1^\star = f' \delta_1 + \Delta \dot{\delta} \ , \qquad \dot{\delta}_2^\star = f' \delta_2 + \frac{f''}{2} \delta_1^2 - \Delta \dot{\delta} \ , \tag{50}$$

and get the same solution $\delta_1 + \delta_2 = \delta_1^\star + \delta_2^\star$. The idea, analogous to the above use of the RG method, is to choose $\Delta \dot{\delta}$ to be 2nd order and absorb any undesirable parts of $\delta_2$. If we choose $\Delta \dot{\delta} = \frac{f''}{2} \delta_1^2 + f'\left(\delta_1^\star - \delta_1\right)$ we have

$$\dot{\delta}_1^\star = f' \delta_1^\star + \frac{f''}{2} \delta_1^{\star 2} \ , \qquad \dot{\delta}_2^\star = f' \delta_2^\star \ , \tag{51}$$

where I have dropped 3rd order terms. Now, if desired, we can set the initial value of $\delta_2^\star$ to zero and forget about it. All we have left is the equation for $\delta_1^\star$. This looks like nothing more than a very long-winded way of saying "if you want to solve $\dot{\delta} = f(\delta)$ to 2nd order, solve the truncated equation $\dot{\delta} = f'\delta + \frac{f''}{2}\delta^2$", and at the level that I will use it here, that is really all it is. However, there was potential for more than that within the argument, in that $\Delta \dot{\delta}$ could have been chosen to be something else if this was convenient to, say, remove a particularly divergent part of $\delta_2$ while still producing an easy to solve equation for $\delta_1^\star$. Also, $\delta_2^\star$ could in principle have been given some part of the initial conditions. I present this mainly as an explanation of the connection between the RG method and the method of simply numerically solving truncated equations. The difference between solving these truncated non-linear equations and the original perturbation equations is that all of the feedback between 2nd and 1st order is included, rather than leaving the first order solution fixed while computing the 2nd order solution.

As suggested in [55], and implemented in [52], one approach to computing the power spectrum of the fields is to use the time evolution equations for the fields to derive time evolution equations for the power spectrum, and to solve those numerically. Unfortunately, these equations involve the bispectrum, which must then be solved for, and the bispectrum evolution equations involve the trispectrum, etc., i.e., one has an infinite hierarchy of equations. The

method of truncating this hierarchy by setting the connected trispectrum to zero is equivalent to applying the above reasoning about truncating the series of perturbative equations, i.e., at linear order in the initial power spectrum the perturbative equations only contain the power spectrum, at 2nd order they only contain the power spectrum and bispectrum, while the trispectrum is needed at 3rd order. This type of approach seems like the only option for including the velocity dispersion as discussed in this paper in high precision calculations, because one cannot solve the evolution equations analytically when including the velocity dispersion. One has to recognize that standard PT could not even produce the analytic results that it does (for time evolution – there are generally numerical integrals over $\mathbf{k}$) without the special fact that, to a very good approximation, one can use the solution to the evolution equations in an Einstein-de Sitter Universe in a realistic Universe as well – this is a very fragile situation and any added complication tends to produce un-solvable equations (e.g., non-negligible massive neutrinos are a good example of this, where, additionally, equations including velocity dispersion would naturally appear explicitly [51, 81]), and this approximation can never describe the kind of dependence on the equation of state of dark energy found by [82].

In anticipation of eventually incorporating the velocity dispersion into a scheme for evolving the power spectrum and higher order statistical equations (the alternative method of closing the hierarchy in [53] may also be a useful way to do this), I write here the evolution equations at 1st order in the power spectrum. Without the new velocity dispersion parts, these would just be the equations for the standard linear theory power spectrum. (I am still assuming an EdS Universe, although that is not necessary here.)

$$P'_{\delta\delta}(k) = 2P_{\delta\theta}(k) \tag{52}$$

$$P'_{\delta\theta}(k) = \frac{3}{2}P_{\delta\delta}(k) - \frac{1}{2}P_{\delta\theta}(k) + P_{\theta\theta}(k) - k^2\tilde{T}\left[P_{\delta\delta}(k) + 2P_{\delta\pi}(k)\right] \tag{53}$$

$$P'_{\theta\theta}(k) = 3P_{\delta\theta}(k) - P_{\theta\theta}(k) - 2k^2\tilde{T}\left[P_{\delta\theta}(k) + 2P_{\theta\pi}(k)\right] \tag{54}$$

$$P'_{\delta\pi}(k) = P_{\delta\theta}(k) + P_{\theta\pi}(k) - \left(1 + \frac{d\ln\tilde{T}}{d\eta}\right)P_{\delta\pi}(k) \tag{55}$$

$$P'_{\theta\pi}(k) = P_{\theta\theta}(k) + \frac{3}{2}P_{\delta\pi}(k) - \left(\frac{3}{2} + \frac{d\ln\tilde{T}}{d\eta}\right)P_{\theta\pi}(k) - k^2\tilde{T}\left[P_{\delta\pi}(k) + 2P_{\pi\pi}(k)\right] \tag{56}$$

$$P'_{\pi\pi}(k) = 2P_{\theta\pi}(k) - 2\left(1 + \frac{d\ln\tilde{T}}{d\eta}\right)P_{\pi\pi}(k) \tag{57}$$

$$\tilde{T}' + \tilde{T} = \frac{2}{3}\,\tilde{T}\,\frac{1}{2\pi^2}\int_0^\infty dq\; q^2 P_{\theta\pi}(q) \tag{58}$$

Intuitive understanding of this last term is that it measures the tendency for velocities to converge on overheated places, and vice versa (remember that, for the definitions we're using, positive $\theta$ and $\pi$ mean, respectively, convergence and, in a one-dimensional sense at least, negative 2nd derivative of the dispersion). Note that, for this paper, writing all of these equations involves some redundancy, in that one could just solve Eqs. (14-16) to get growth factors as functions of $k$, using them to evaluate Eq. (58). When one goes to higher order, however, where $\delta$, $\theta$, and $\pi$ are no longer perfectly correlated, one will need all the equations.

### 2. Power law power spectra

Before I solve these equations completely numerically, it is informative to assume a power law power spectrum, for which we can derive some scalings. Defining $k_{\rm NL}$ by $\Delta_{\rm L}^2(k_{\rm NL}) = 1$, and using time evolution $\Delta_L(k) \propto D^2$, I find $k_{\rm NL} \propto D^{-2/(3+n)}$. The velocity dispersion must follow

$$\tilde{T} \propto k_{\rm NL}^{-2} \propto D^{\frac{4}{3+n}}\;. \tag{59}$$
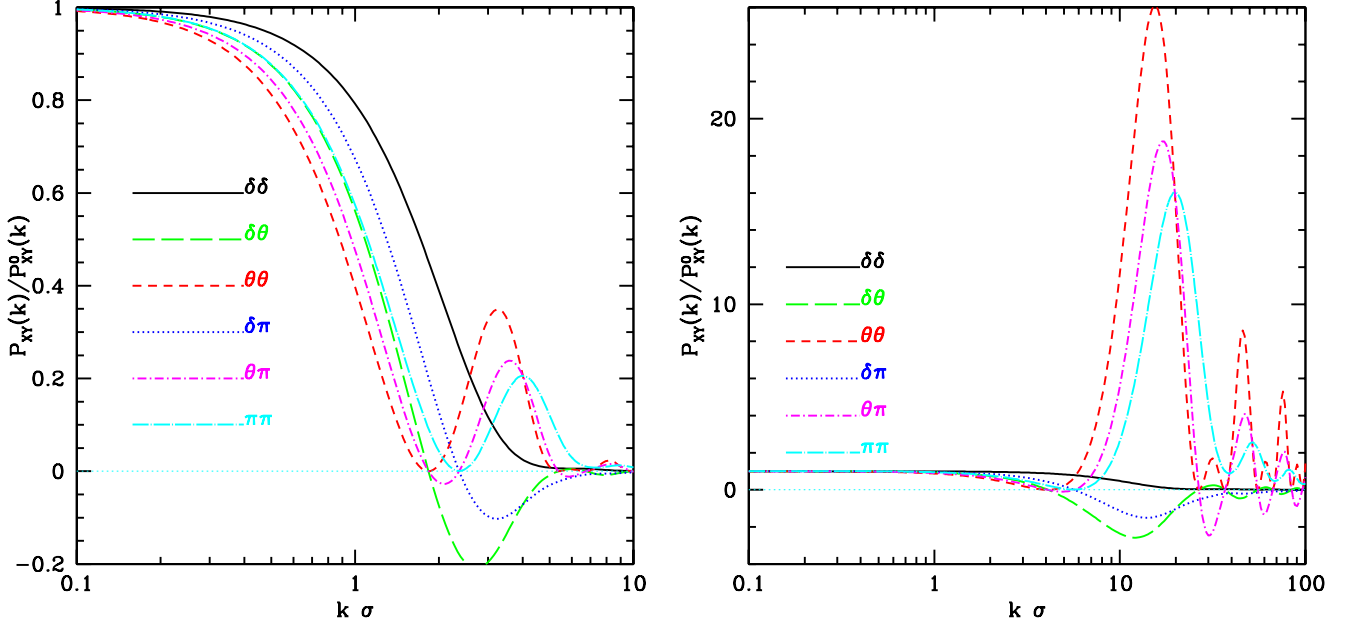
FIG. 1: Ratio of linear power with Jeans-like suppression by velocity dispersion to without (i.e., to the $\tilde{T} \to 0$ limit). The left panel is for a power law power spectrum with $n = -1.4$, while the right is for $n = -2.75$. Lines are identified in the figure panel. $\sigma \equiv \tilde{T}^{1/2}$.

Inevitably, this is the same evolution that the RG calculation above predicted. For this to be consistent with Eq. (58), we must have

$$\langle \theta \pi \rangle = \frac{6}{3+n} + \frac{3}{2} \; , \tag{60}$$

i.e., as structure grows on large scales, the Jeans-like effect of velocity dispersion must erase enough small-scale structure to maintain fixed $\langle \theta \pi \rangle$, with the fixed value being larger for power spectra where there is less small-scale power. Again, this result is identical to the RG result. This applies only for $n > -3$ of course – for $n < -3$ this discussion breaks down right at the beginning, with the definition of $k_{\rm NL}$. We can also solve analytically for the growing mode amplitude of $\pi$ in the $k \to 0$ limit, which is suppressed relative to $\delta$ and $\theta$ because of the $d \ln \tilde{T}/d\eta$ factor (this came from the fact that $\pi$ fluctuations are measured relative to the changing background dispersion). The result, again the same as in the RG calculation, is

$$\pi(k \to 0) = \frac{1}{2} \frac{3+n}{5+n} \delta \, (k \to 0) \; , \tag{61}$$

i.e., the dispersion fluctuations are smaller for models with less small-scale power, although the best way of looking at this is probably to think of these models as the ones where the non-linear scale is changing more quickly, and therefor the homogeneous dispersion is increasing more quickly, diluting the perturbations. All that remains to be calculated (for a power law power spectrum) are the high-$k$ suppression kernels that must be applied to each field, which generally must be obtained numerically.

Figure 1 shows these kernels, i.e., the effect of the Jeans-like smoothing due to the mean velocity dispersion, for power law power spectra with $n = -1.4$, corresponding to roughly the present non-linear scale, and $-2.75$, appropriate to the non-linear scale at a much earlier time. To be clear, Fig. 1 was made by evolving Eqs. (52-58) numerically, with initial conditions for a scaling solution determined by first evolving assuming the scaling solution for $\tilde{T}$, then restarting using the results for the initial kernels and normalization (at that point, the system will evolve stably on the scaling solution). Generally, the initial conditions would break the scaling, e.g., if one starts with the bare un-smoothed power law, that is not consistent with scaling (one would think that non-linear effects would erase the memory of a break in scaling in the distant past, but that does not happen at linear order). In each case Fig. 1 shows the power spectrum from the numerical calculation divided by the power spectrum that would be obtained by taking the $\tilde{T} \to 0$ limit (specifically, taking the overall normalization of $\tilde{T}$ to zero, which does not affect $d \ln \tilde{T}/d\eta$ which sets

the normalization of $\pi$). The kernels for the different power laws are dramatically different when plotted as functions of $k\sigma$ ($\sigma \equiv \tilde{T}^{1/2}$). Additionally, the relation between the non-linear scale and $\sigma$ is much different: $\sigma = 0.19k_{\rm NL}^{-1}$ for $n = -1.4$, while $\sigma = 0.000023k_{\rm NL}^{-1}$ for $n = -2.75$, i.e., the dispersion scale lags much farther behind the non-linear scale for the spectrum with less small-scale power, not surprisingly. Note that the non-linear and dispersion scales are increasing extremely rapidly for $n = -2.75$, like $a^8$, so that the dispersion scale only lags the non-linear scale by a factor of $\sim 1/4$ in expansion factor. Similarly, even though the extreme-looking spike in the kernel covers a factor $\sim 2$ in $k$, a given mode only spends $\sim 10\%$ of an expansion factor within this feature. In spite of the apparently large differences, there is a remarkable relation between the two power law cases: the expansion factor by which the dispersion scale lags the non-linear scale, i.e., $(\sigma\, k_{\rm NL})^{(3+n)/2}$, is essentially identical between the two cases – 0.262 vs. 0.265 from my numerical solution, for $n = -1.4$ and $n = -2.75$. To be clear, the dispersion in each case is determined by the requirement that Eq. (60) for $\langle\theta\pi\rangle$ is satisfied after the power is filtered by the complex-looking kernels in Fig. 1. Eq. (45) actually does a reasonable job anticipating this relation, in spite of its approximations, predicting 0.18 and 0.12 for $n = -1.4$ and $n = -2.75$, respectively, but it is not clear if there is any deep reason for the near perfect agreement in the exact calculation (the under-prediction by Eq. 45 is understandable, as the Gaussian approximation I used there for the kernels works essentially perfectly for the initial suppression, but misses the wiggles, which produce extra dispersion).

### 3. $\Lambda CDM$

The situation for a $\Lambda$CDM power spectrum is less straightforward than for power law power spectra, where we could imagine the initial dispersion was arbitrarily small, yet still obtain an interesting dispersion tied to the non-linear scale. In the real Universe, we have specific physically motivated initial conditions, and we cannot ignore them (at least not within the lowest order calculation in this paper – as I discuss below, I actually expect that the influence of the smallest scale initial conditions will be erased when the calculation is taken to the next order). A typical WIMP with mass $m \sim 100$ GeV decouples thermally from the relativistic electrons, positrons, and neutrinos at $T_d \sim 20$ MeV [83]. The coupling to the relativistic plasma actually produces acoustic oscillations in the transfer function, which I will ignore. The 1D velocity dispersion at kinetic decoupling is $v_d^{\rm 1D} = \left(\frac{2}{3}\frac{T_d}{m}\right)^{1/2} c$. For a typical WIMP, the scale below which power is erased is of order the scale of free-streaming after decoupling, but would actually be similar even for much more massive (i.e., slow moving) particles, because a substantial fraction of the effect is due to Silk-damping-like friction between the WIMPs and leptons during decoupling [83]. Using Eq. (48) of [83], I will assume the power entering the matter dominated era is suppressed by $\exp(-k^2/k_{fs}^2)$ with $k_{fs}^{-1} \sim \left(\frac{6}{5}\frac{T_d}{m}\right)^{1/2} \tau_d \ln\left(\tau_{eq}/1.05\tau_d\right)$ where $\tau_d$ is the conformal time at decoupling and $\tau_{eq}$ is the conformal time at matter-radiation equality (after which the free streaming has frozen). For the parameters I am using, the primordial velocity dispersion, in Hubble flow distance units, extrapolated to the present time assuming no enhancement by structure formation, would be $\tilde{T}^{1/2} = 8.6 \times 10^{-10}\, h^{-1}$ Mpc, and the smoothing scale is $k_{fs}^{-1} = 10^{-6}\, h^{-1}$ Mpc.

The total density variance in a realistic $\Lambda$CDM model, in linear theory, without any enhancement of velocity dispersion by structure formation, is $\langle\delta_1^2\rangle \simeq 517D^2$, where $D$ is the linear theory growth factor normalized to be 1 at the present time. If we could assume that $\sigma_{\theta\pi}^2 = \langle\theta_1\pi_1\rangle$ followed the same form, as it does before there is any feedback (and assume an EdS Universe), the solution to Eq. (58) would be $\tilde{T}/\tilde{T}_i = \exp\left[517\left(a^2 - a_i^2\right)/3 - \ln\left(a/a_i\right)\right]$. We see that if $\langle\theta_1\pi_1\rangle$, which starts out equal to $\langle\delta_1^2\rangle$, remains anywhere near as large, the current dispersion would be extremely large, i.e., $\tilde{T}/\tilde{T}_i \sim \exp(517/3) \sim 10^{75}$. Clearly this is too large, which is primarily explained by the fact that, once the velocity dispersion begins to increase, $\langle\theta_1\pi_1\rangle$ falls well below $\langle\delta_1^2\rangle$, as we saw for power laws in Eq. (61). We can estimate the size of this effect by continuing to ignore any additional smoothing by the extra dispersion, but solving Eqs. (60) and (61) to find an effective $n$ that will allow both equations to be satisfied, given the large density variance $\langle\delta_1^2\rangle$. The result is $\langle\theta_1\pi_1\rangle \simeq \left(\frac{3}{2}\langle\delta_1^2\rangle\right)^{1/2}$, or, for $\langle\delta_1^2\rangle = 517D^2$, $n = -2.77$ (at $D = 1$) and $\langle\theta_1\pi_1\rangle = 27.1D$. Using this we can solve Eq. (58) to obtain $\tilde{T}/\tilde{T}_i \sim \exp[(2/3)27] \sim 10^8$. This may seem like a big number, but it is in fact inadequate to produce a dispersion that is dynamically significant at late times, i.e., we need a boost by $\sim 9$ orders of magnitude in $\tilde{T}^{1/2}$, not just 4. A full numerical evaluation of the evolution equations for $\tilde{T}$ and the power spectra basically confirms this analysis. $\tilde{T}^{1/2} = 6 \times 10^{-6}\, h^{-1}$ Mpc at the present time, $\sim 4$ orders of magnitude larger than the primordial dispersion but nowhere near macroscopic, and not even large enough to begin suppressing the power much beyond the primordial $k_{fs}^{-1} = 10^{-6}\, h^{-1}$ Mpc. The full calculation gives $\langle\delta_1^2\rangle = 506$ and $\langle\theta_1\pi_1\rangle = 26.4$, close to the no-smoothing approximation.

Now, the derivation of Eq. (58) did not assume linear theory, only that velocities are irrotational, and that we can drop higher moments of the velocity distribution function, so it would be correct to use a non-linear version of $\langle\theta\pi\rangle$ to evaluate it. For example, the standard 1-loop PT power spectrum has $\langle\delta^2\rangle > 10^5$, easily large enough to

completely change the picture for the growth of dispersion, i.e., the feedback between the production of dispersion by non-linearity and the smoothing of the perturbations by dispersion will be critical. Exactly how this works out is impossible to say at this point. We don't have higher order PT calculations that include this form of velocity dispersion, and, as I found above even at linear order, it is almost surely impossible to derive analytic PT results with dispersion. The next step appears to be to set up and solve numerically the non-linear versions of Eqs. (52-57), and the inevitably accompanying bispectrum equations (at least). It seems very likely that this will produce the level of dispersion necessary to produce the percent level corrections to the power spectrum at $k \sim 0.1\,h\,\mathrm{Mpc}^{-1}$ found by [67]. This requires a dispersion level corresponding to $\sim 1\,h^{-1}\,\mathrm{Mpc}$. For non-linear scale $k_{\mathrm{NL}}$, and $n \sim -1.4$, the result I found for a power law power spectrum was just this $\sim 1\,h^{-1}\,\mathrm{Mpc}$. This would also be consistent with the scale at which the RG evolution in [55] saturated to it's fixed-point slope of $n = -1.4$.

## III.   REDSHIFT-SPACE DISTORTIONS

Interpretation of galaxy redshift surveys generally requires us to account for the distortion of the observed clustering due to Doppler shifts caused by peculiar velocities [30, 84, 85, 86, 87], including, potentially, those due to velocity dispersion as well as bulk flows. Here I will continue to discuss dark matter only – the differences between galaxies and dark matter, while certainly important, are the subject for another paper.

Starting from the distribution function, defined above, it is straightforward to derive the perturbative redshift-space distortions including possibly velocity dispersion. Defining the redshift-space coordinate for a particle, $\mathbf{s} = \mathbf{x} + \hat{r}\,p_{\parallel}/am\mathcal{H}$ where $\hat{r}$ is the unit vector pointing along the observer's line of sight and $p_{\parallel} = \mathbf{p} \cdot \hat{r}$, the density in redshift space is:

$$\rho_s\left(\mathbf{s}\right) = m\,a^{-3} \int d^3\mathbf{x}\,d^3\mathbf{p}\,f\left(\mathbf{x},\mathbf{p}\right)\delta^D\left(\mathbf{s} - \mathbf{x} - \hat{r}\frac{p_{\parallel}}{am\mathcal{H}}\right) = m\,a^{-3}\int d^3\mathbf{p}\,f\left(\mathbf{s} - \hat{r}\frac{p_{\parallel}}{am\mathcal{H}},\mathbf{p}\right)\ . \tag{62}$$

I now expand the distribution function in a Taylor series in $p_{\parallel}/am\mathcal{H}$,

$$\rho_s\left(\mathbf{s}\right) = m\,a^{-3}\int d^3\mathbf{p} f\left(\mathbf{s},\mathbf{p}\right) + m\,a^{-3}\int d^3\mathbf{p}\frac{\partial f\left(\mathbf{s},\mathbf{p}\right)}{\partial s_{\parallel}}\left(-\frac{p_{\parallel}}{am\mathcal{H}}\right) + \frac{1}{2}m\,a^{-3}\int d^3\mathbf{p}\frac{\partial^2 f\left(\mathbf{s},\mathbf{p}\right)}{\partial s_{\parallel}^2}\left(\frac{p_{\parallel}}{am\mathcal{H}}\right)^2 + \ldots \tag{63}$$

Taking the $s_{\parallel}$ derivatives outside of the $\mathbf{p}$ integrals, and using the above relations between moments of the distribution function with respect to momentum and bulk velocity and velocity dispersion, I obtain

$$\delta_s\left(\mathbf{s}\right) = \delta\left(\mathbf{s}\right) - \frac{1}{\mathcal{H}}\frac{\partial}{\partial s_{\parallel}}\left[\left(1+\delta\right)v_{\parallel}\right] + \frac{1}{2\mathcal{H}^2}\frac{\partial^2}{\partial s_{\parallel}^2}\left[\left(1+\delta\right)\left(v_{\parallel}^2 + \sigma_{\parallel\parallel}\right)\right] + \ldots \tag{64}$$

The linearized version of this equation, with zero-order velocity dispersion $\bar{\sigma}_{\parallel\parallel} = \mathcal{H}^2\tilde{T}_0$, is

$$\delta_s\left(\mathbf{s}\right) = \delta\left(\mathbf{s}\right) - \frac{1}{\mathcal{H}}\frac{\partial v_{\parallel}}{\partial s_{\parallel}} + \frac{\tilde{T}_0}{2}\frac{\partial^2\delta}{\partial s_{\parallel}^2} + \frac{1}{2\mathcal{H}^2}\frac{\partial^2\sigma_{\parallel\parallel}}{\partial s_{\parallel}^2} \tag{65}$$

The first two terms are the usual linear theory result [30], while the last two are related to the inclusion of dispersion. In Fourier space, and using $v^i = -i\mathcal{H}k^i k^{-2}\theta$ and $\sigma_{ij} = 2\mathcal{H}^2\tilde{T}_0 k_i k_j k^{-2}\pi$ (for irrotational velocities), this is

$$\delta_s\left(\mathbf{k}\right) = \left(1 - \frac{1}{2}k^2\tilde{T}_0\,\mu^2\right)\delta\left(\mathbf{k}\right) + \mu^2\,\theta\left(\mathbf{k}\right) - k^2\tilde{T}_0\,\mu^4\,\pi\left(\mathbf{k}\right)\ . \tag{66}$$

where I have not used the usual $\theta = f\delta = \frac{d\ln D}{d\ln a}\delta$ simplification because the relation between $\theta$ and $\delta$ is complicated by the presence of dispersion (their Jeans smoothing kernels are different – note that, like [67], I found in Fig. 1 that the velocity divergence is suppressed substantially more than the density fluctuations, which means that these corrections will be more important for redshift space distortions than they would be in real space). Similarly, $\pi$ will only be proportional, not equal, to $\delta$ on large scales (usually smaller), and will have a different smoothing kernel. The first dispersion term looks like the first term in a small-$k$ expansion of the smoothing kernels that have been used to model non-linear peculiar velocities in the past [88]. There will clearly be a series of similar higher order terms that could be re-summed into a simple smoothing kernel. Note, however, that there will also be a series of terms involving the homogeneous kurtosis of the velocity distribution, which generally could be of similar order to the re-summed terms, so it isn't obvious that the re-summation will be particularly useful (i.e., once the re-summed terms are needed, other

independent terms may be needed as well). Also, note that in cases where the velocity dispersion is relatively small, the similar distortion coming from the $v_\parallel^2$ term in Eq. (64) could easily dominate over this term.

Clearly, redshift-space distortions will require further consideration as the higher order program suggested by this paper is carried out. The suppression of the new terms by $k^2 \tilde{T}_0$ probably means that they will not be significant unless terms higher order in $\delta$ are also significant on the same scale, i.e., $k^2 \tilde{T}_0$ probably should be considered to be a new expansion parameter, at least as small as $\delta$, and probably $\delta^2$ (this is consistent with the way that $\tilde{T}_0$ is derived, as proportional to the linear power spectrum, however, the correspondence is not unambiguous because of the very different origin and form of growth of the terms). In any case, interpretation of very high precision observations, on the imperfectly linear scales where most of the information resides, will clearly require a more subtle calculation than has so far been done, however, one should not despair, or flee to relying entirely on numerical simulations, because an understanding of the detailed redshift-space structure of observations will ultimately make our measurements of fundamental physics much more robust and believable. Note that nothing requires, or even really suggests, that the effective $\tilde{T}_0$ for galaxies [89, 90, 91, 92, 93] must be equal to that for dark matter – in the end, it will probably be a completely free parameter for galaxies, like bias.

## IV. DISCUSSION

The deficiencies of standard perturbation theory are a lack of control of the higher moments of the velocity distribution function, beyond the pressureless fluid approximation, problems with accuracy related to the fact that PT integrals include small scales that are generally highly non-linear, even when the scales we are interested in are weakly non-linear, and, of course, simply not working on small scales where the fluctuations are large. The deficiencies of numerical simulations are speed, and the related fact that statistics (e.g., the power spectrum) cannot be computed directly, but only as averages over realizations of the random density field, which must contain many orders of magnitude more degrees of freedom (e.g., a billion) than one really needs to describe the statistic of interest (e.g., a few dozen for the power spectrum). Additionally, and probably most importantly, the cumbersome, opaque nature of simulation results is greatly exacerbated when they are used to model galaxies or other observables instead of just dark matter, while the advantage of being more or less exact for gravity alone (at least in the straightforward limits of large box size and particle density) is lost.

This paper directly addresses the PT deficiency of missing higher moments of the velocity distribution function, showing how they can be re-activated and generated at a significant level, starting from first principles. The approach here may also improve PT by providing natural regulation of small-scale structure, i.e., the Jeans smoothing that arises here has the same kind of effect as the propagator resummation of [47, 48]. The philosophy of this paper is that the cold "streams" often discussed as "crossing" are mythical objects – what one sees in the real Universe is always some evolution of effectively warm (although maybe not very warm) material. The meaning of this idea is very clear for power law power spectra, where, as we saw, non-trivial effects of velocity dispersion can be computed without any initial dispersion entering the discussion. The dispersion bootstraps itself up from an arbitrarily small start. The temperature is locked into a sort of equilibrium with the growth of structure. The situation is not as clear for $\Lambda$CDM power spectra, not so much because the $\Lambda$CDM is cold as because there is very little small-scale power in these models, so the dispersion computed to linear order in the power spectrum remains well below the non-linear scale (although orders of magnitude larger than it would be if there was no structure). This situation will change when calculations are done to higher order, where I showed that there is enough small-scale power generated to produce dispersion that would be far too large in absence of feedback on the structure formation itself. The initial velocity dispersion should be rendered irrelevant when the system moves into a sort of self-regulating mode, like the power law example, where the velocity dispersion and growth of structure are tightly coupled by feedback between them. Ultimately, it seems likely that the best effective small-scale model for doing perturbation theory will involve some more general balance between different pieces, i.e., density, velocity divergence, dispersion, and maybe even vorticity, etc., because we know that physically this is what the small-scale structure really is, i.e., halos which can only be described as a delicate balance of the original perturbative LSS quantities. To put this another way: hopefully, when all of the relevant elements are included, there will be some fixed-point structure for small scales with a clear physical interpretation and effectively far fewer than the original number of degrees of freedom (akin to the fixed-point power law found in [55], but with richer structure).

This paper does not exactly contain useful quantitative take-home results. The results are primarily a procedure to follow for future calculations. If there is a single equation that best represents the results, it is probably Eq. (45), which shows how, for a power law power spectrum, the velocity dispersion tracks the non-linear scale, with Jeans filtering erasing more and more small-scale structure as the larger scale structure grows. Eqs. (46) and (58) are also critical, in that they show generally how homogeneous velocity dispersion is generated out of fluctuations. The key new variable, equivalent to $\delta$ and $\theta$, but for perturbations in velocity dispersion, is $\pi \propto \partial_i \partial_j \sigma^{ij}$.

The next step in this line of work is to derive to the next-order equations like Eq. (52-57) (but including bispectrum equations), which are needed to have any chance of properly describing $\Lambda$CDM. Then the results can be tested by comparison to numerical simulations. While $\Lambda$CDM is of course the ultimate goal, tests of the theoretical concepts here might be more conveniently done with power law simulations, particularly ones with relatively blue spectra, as in [54]. Any of the renormalization approaches that have been developed recently can probably be applied to the problem of renormalizing the velocity dispersion, at least in principle, i.e., resumming a set of terms that generates the dispersion, either explicitly or through a renormalization group equation. The obstacle to doing this elegantly may be the difficulty of obtaining analytic solutions as the evolution equations become more complicated (this problem of requiring analytic solutions is, I think, the primary reason to favor the "numerical evaluation of evolution equations for statistics" approach advocated in this paper over other, more completely analytic, recent approaches).

One might ask "why stop with the 2nd moment of the distribution function, i.e., why can we drop $q_{ijk}$ from Eq. (8)?" One hope, which will need to be verified by future calculations, is that the effect of increasingly high moments on large scales may take the form of a gradient expansion, i.e., in Fourier space a series where the effects of increasingly high moments enter multiplied by increasing powers of $k$. In this case, we would expect that, on scales where the effect of the 2nd moment are already small, the effects of higher moments will be even smaller.

The result for redshift-space space distortions (Eq. 64) leads to the question: Why do we use $\mathbf{v}$, and in this paper $\sigma_{ij}$, as the variables to be solved for perturbatively, rather than, e.g., momentum $(1+\delta)\mathbf{v}$ and kinetic energy $(1+\delta)(v_i v_j + \sigma_{ij})$? An answer one might have considered was that velocities are needed to compute redshift-space distortions, but here we see that the most direct quantities for that are in fact momentum and energy, with additional perturbative calculations needed when starting from $\delta$, $\mathbf{v}$, and $\sigma_{ij}$. A change to total kinetic energy would have the potentially substantial effect of increasing the Jeans-like smoothing of the linear power, because the zero order pressure would be larger. I do not see any clear a priori reason to favor one option over another – they are just different ways of arranging a series of terms, which should be equivalent if one could include an infinite number of terms. Note that, while the choice of variables is optional, the renormalization of the energy-related variable is not optional – it simply makes no sense to ignore the fact that the size of terms in the series describing one of your variables is increasing rapidly, rather than decreasing, with order. There is a lot of circumstantial evidence that using total kinetic energy could be useful. The renormalization/resummation of the propagator in [47] leads to filtering much like the Jeans filtering we find due to velocity dispersion, but with scale given by the bulk velocity power spectrum. The Lagrangian PT-based scheme of [49] leads to a similar result. Another possibility to consider would be the evolution of large-scale fields with the small-scale structure explicitly averaged out, which would naturally lead to the inclusion of small-scale bulk velocities in the dispersion term [94, 95, 96]. One might also consider using the Schrödinger equation representation of the exact Vlasov equation, proposed in [97], combined with the approach of this paper.

In the end, one could view the approach in this paper as a re-activation of the program under discussion in papers like [98], which set out to integrate the BBGKY equations numerically. This reconsideration is timely because of several developments over the last thirty years. We now know the appropriate class of models to focus on, especially including the appropriate the initial conditions. There has been a lot of work on both perturbation theory beyond linear order, and on N-body simulations, with the limitations of each teaching us a lot about what is needed from new methods. We also have a specific calculation to focus on: the scales where baryonic acoustic oscillations (BAO) are observable [19, 99, 100, 101, 102, 103, 104, 105, 106, 107], which points us toward the perturbative approach that motivates truncating the hierarchy and believing that high precision can be achieved (in contrast to [98], who were focused on the more strongly non-linear regime, where the truncation used here is not well-motivated). The same scales are also potentially the most powerful for other measurements based on, e.g., redshift-space distortions [3].

The basic form of calculation I do here is completely standard in some other areas of cosmology. For example, the evolution of the homogeneous (background) value of an interacting scalar field in the early Universe is affected by quantum and thermal fluctuations. Its evolution is not described by the original equation of motion with all perturbations ignored, but by a renormalized effective potential, which is at least formally infinitely different from the original potential, i.e., completely dominated by the part due to fluctuations [108]. Another interestingly similar calculation is the development of equilibrium after preheating after inflation. [109, 110, 111, 112, 113] perform fully non-linear simulations of the interaction of scalar fields, much like large-scale structure simulations, with the added twist that the background density and pressure are affected by the fluctuations. The evolution of the scale factor is calculated by taking spatial averages over the fluctuations as the simulation is running. In the beginning, the nearly homogeneous inflaton field dominates, and a very naive calculation might compute the expansion of the simulation in advance assuming homogeneous evolution, but by the end of the simulation the homogeneous component has actually practically disappeared, with the background evolution completely dominated by the fluctuations, which behave like radiation. Of course no one would ever do the very naive calculation just mentioned, where the affect of the fluctuations on the background equation of state is ignored... except that this is what we do when we assume that the tiny initial temperature of CDM means that it will forever remain cold.

Substantial work remains to determine if the approach in this paper enhances the accuracy of predictions of quasi-linear clustering in the real Universe; however, it is now possible to consider an effect that was previously outside of any first-principles computational control in this kind of perturbation theory. The bottom line of this paper is that even linear theory fluctuations necessarily imply a significant one-loop renormalization of the background velocity dispersion.

[1] U. Seljak, A. Slosar, and P. McDonald, Journal of Cosmology and Astro-Particle Physics **10**, 14 (2006).
[2] U. Seljak, A. Makarov, P. McDonald, and H. Trac, Physical Review Letters **97**, 191303 (2006).
[3] P. McDonald and U. Seljak, ArXiv e-prints (2008), 0810.0323.
[4] P. McDonald, Phys. Rev. D **78**, 123519 (2008), 0806.1061.
[5] P. McDonald, ArXiv e-prints (2009), 0907.5220.
[6] R. Scoccimarro, ArXiv e-prints (2009), 0906.4545.
[7] K. C. Chan and R. Scoccimarro, ArXiv e-prints (2009), 0906.4548.
[8] E. Komatsu, J. Dunkley, M. R. Nolta, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. Limon, L. Page, et al., ApJS **180**, 330 (2009), 0803.0547.
[9] U. Seljak, A. Makarov, P. McDonald, S. F. Anderson, N. A. Bahcall, J. Brinkmann, S. Burles, R. Cen, M. Doi, J. E. Gunn, et al., Phys. Rev. D **71**, 103515 (2005).
[10] E. Sefusatti and E. Komatsu, Phys. Rev. D **76**, 083004 (2007), arXiv:0705.0343.
[11] N. Afshordi, P. McDonald, and D. N. Spergel, ApJ **594**, L71 (2003).
[12] V. Desjacques and U. Seljak, ArXiv e-prints (2009), 0907.2257.
[13] T. Y. Lam, V. Desjacques, and R. K. Sheth, ArXiv e-prints (2009), 0908.2285.
[14] M. Grossi, L. Verde, C. Carbone, K. Dolag, E. Branchini, F. Iannuzzi, S. Matarrese, and L. Moscardini, MNRAS **398**, 321 (2009), 0902.2013.
[15] N. Barnaby and Z. Huang, ArXiv e-prints (2009), 0909.0751.
[16] A. Taruya, K. Koyama, and T. Matsubara, Phys. Rev. D **78**, 123534 (2008), 0808.4085.
[17] M. R. Nolta, J. Dunkley, R. S. Hill, G. Hinshaw, E. Komatsu, D. Larson, L. Page, D. N. Spergel, C. L. Bennett, B. Gold, et al., ApJS **180**, 296 (2009), 0803.0593.
[18] M. Tegmark, D. J. Eisenstein, M. A. Strauss, D. H. Weinberg, M. R. Blanton, J. A. Frieman, M. Fukugita, J. E. Gunn, A. J. S. Hamilton, G. R. Knapp, et al., Phys. Rev. D **74**, 123507 (2006), arXiv:astro-ph/0608632.
[19] P. McDonald and D. J. Eisenstein, Phys. Rev. D **76**, 063009 (2007), arXiv:astro-ph/0607122.
[20] P. McDonald, U. Seljak, R. Cen, D. Shih, D. H. Weinberg, S. Burles, D. P. Schneider, D. J. Schlegel, N. A. Bahcall, J. W. Briggs, et al., Astrophys. J. **635**, 761 (2005).
[21] M. Viel and M. G. Haehnelt, MNRAS **365**, 231 (2006).
[22] P. McDonald, U. Seljak, S. Burles, D. J. Schlegel, D. H. Weinberg, R. Cen, D. Shih, J. Schaye, D. P. Schneider, N. A. Bahcall, et al., ApJS **163**, 80 (2006), arXiv:astro-ph/0405013.
[23] P. McDonald, Astrophys. J. **585**, 34 (2003), arXiv:astro-ph/0108064.
[24] P. McDonald, J. Miralda-Escudé, and R. Cen, Astrophys. J. **580**, 42 (2002).
[25] H. Hoekstra, Y. Mellier, L. van Waerbeke, E. Semboloni, L. Fu, M. J. Hudson, L. C. Parker, I. Tereno, and K. Benabed, Astrophys. J. **647**, 116 (2006), astro-ph/0511089.
[26] N. Aghanim, S. Majumdar, and J. Silk, Reports on Progress in Physics **71**, 066902 (2008), 0711.0518.
[27] T.-C. Chang, U.-L. Pen, J. B. Peterson, and P. McDonald, Physical Review Letters **100**, 091303 (2008), 0709.3672.
[28] U. Seljak and M. Zaldarriaga, Astrophys. J. **469**, 437 (1996).
[29] P. J. E. Peebles, *The large-scale structure of the universe* (Research supported by the National Science Foundation. Princeton, N.J., Princeton University Press, 1980. 435 p., 1980).
[30] N. Kaiser, MNRAS **227**, 1 (1987).
[31] N. Kaiser, ApJ **284**, L9 (1984).
[32] J. Hamann, S. Hannestad, A. Melchiorri, and Y. Y. Y. Wong, Journal of Cosmology and Astro-Particle Physics **7**, 17 (2008), 0804.1789.
[33] A. G. Sánchez and S. Cole, MNRAS **385**, 830 (2008), arXiv:0708.1517.
[34] K. Heitmann, Z. Lukić, P. Fasel, S. Habib, M. S. Warren, M. White, J. Ahrens, L. Ankeny, R. Armstrong, B. O'Shea, et al., Computational Science and Discovery **1**, 015003 (2008), 0706.1270.
[35] K. Heitmann, M. White, C. Wagner, S. Habib, and D. Higdon, ArXiv e-prints (2008), 0812.1052.
[36] K. Heitmann, D. Higdon, M. White, S. Habib, B. J. Williams, and C. Wagner, ArXiv e-prints (2009), 0902.0429.

[37] P. McDonald and A. Roy, Journal of Cosmology and Astro-Particle Physics **8**, 20 (2009), 0902.0991.
[38] R. Juszkiewicz, MNRAS **197**, 931 (1981).
[39] E. T. Vishniac, MNRAS **203**, 345 (1983).
[40] J. N. Fry, Astrophys. J. **279**, 499 (1984).
[41] M. H. Goroff, B. Grinstein, S.-J. Rey, and M. B. Wise, Astrophys. J. **311**, 6 (1986).
[42] R. Scoccimarro and J. A. Frieman, Astrophys. J. **473**, 620 (1996), arXiv:astro-ph/9602070.
[43] F. Bernardeau, S. Colombi, E. Gaztanaga, and R. Scoccimarro, Physics Reports **367**, 1 (2002).
[44] D. Jeong and E. Komatsu, Astrophys. J. **651**, 619 (2006), astro-ph/0604075.
[45] S. Dürr, Z. Fodor, J. Frison, C. Hoelbling, R. Hoffmann, S. D. Katz, S. Krieg, T. Kurth, L. Lellouch, T. Lippert, et al., Science **322**, 1224 (2008), 0906.3599.
[46] W.-M. Yao and et al., Journal of Physics G Nuclear Physics **33**, 1 (2006), arXiv:astro-ph/0601168.
[47] M. Crocce and R. Scoccimarro, Phys. Rev. D **73**, 063520 (2006).
[48] M. Crocce and R. Scoccimarro, Phys. Rev. D **73**, 063519 (2006).
[49] T. Matsubara, Phys. Rev. D **77**, 063530 (2008), arXiv:0711.2521.
[50] M. Crocce and R. Scoccimarro, Phys. Rev. D **77**, 023533 (2008), arXiv:0704.2783.
[51] J. Lesgourgues, S. Matarrese, M. Pietroni, and A. Riotto, Journal of Cosmology and Astro-Particle Physics **6**, 17 (2009), 0901.4550.
[52] M. Pietroni, Journal of Cosmology and Astro-Particle Physics **10**, 36 (2008), 0806.0971.
[53] A. Taruya and T. Hiramatsu, Astrophys. J. **674**, 617 (2008), 0708.1367.
[54] L. M. Widrow, P. J. Elahi, R. J. Thacker, M. Richardson, and E. Scannapieco, MNRAS **397**, 1275 (2009), 0901.4576.
[55] P. McDonald, Phys. Rev. D **75**, 043514 (2007).
[56] J. Carlson, M. White, and N. Padmanabhan, Phys. Rev. D **80**, 043531 (2009), 0905.0479.
[57] T. Nishimichi, A. Shirata, A. Taruya, K. Yahata, S. Saito, Y. Suto, R. Takahashi, N. Yoshida, T. Matsubara, N. Sugiyama, et al., PASJ **61**, 321 (2009), 0810.0813.
[58] F. Bernardeau, M. Crocce, and R. Scoccimarro, Phys. Rev. D **78**, 103521 (2008), 0806.2334.
[59] F. Bernardeau and P. Valageas, Phys. Rev. D **78**, 083503 (2008), 0805.0805.
[60] S. Matarrese and M. Pietroni, Journal of Cosmology and Astro-Particle Physics **6**, 26 (2007), arXiv:astro-ph/0703563.
[61] S. Matarrese and M. Pietroni, Modern Physics Letters A **23**, 25 (2008), arXiv:astro-ph/0702653.
[62] P. Valageas, A&A **484**, 79 (2008), 0711.3407.
[63] K. Izumi and J. Soda, Phys. Rev. D **76**, 083517 (2007), arXiv:0706.1604.
[64] T. Matsubara, Phys. Rev. D **78**, 083519 (2008).
[65] P. McDonald, Phys. Rev. D **74**, 103512 (2006).
[66] U. Seljak, A. Makarov, R. Mandelbaum, C. M. Hirata, N. Padmanabhan, P. McDonald, M. R. Blanton, M. Tegmark, N. A. Bahcall, and J. Brinkmann, Phys. Rev. D **71**, 043511 (2005).
[67] S. Pueblas and R. Scoccimarro, Phys. Rev. D **80**, 043504 (2009), 0809.4606.
[68] P. Valageas, A&A **379**, 8 (2001).
[69] M. Shoji and E. Komatsu, Astrophys. J. **700**, 705 (2009), 0903.2669.
[70] J. Gaite and A. Domínguez, Journal of Physics A Mathematical General **40**, 6849 (2007), arXiv:astro-ph/0610886.
[71] T. Tatekawa, Phys. Rev. D **70**, 064010 (2004), arXiv:astro-ph/0408120.
[72] H. Sotani and T. Tatekawa, Phys. Rev. D **73**, 024024 (2006), arXiv:astro-ph/0601281.
[73] A. L. Ribeiro and J. G. Peixoto de Faria, Phys. Rev. D **71**, 067302 (2005), arXiv:astro-ph/0502580.
[74] N. Menci, MNRAS **330**, 907 (2002), arXiv:astro-ph/0111228.
[75] T. Buchert and A. Domínguez, A&A **438**, 443 (2005), arXiv:astro-ph/0502318.
[76] L.-Y. Chen, N. Goldenfeld, and Y. Oono, Physical Review Letters **73**, 1311 (1994).
[77] T. Kunihiro, Progress of Theoretical Physics **94**, 503 (1995), hep-th/9505166.
[78] T. Kunihiro, Progress of Theoretical Physics **97**, 179 (1997), hep-th/9609045.
[79] D. V. Shirkov and V. F. Kovalev, Physics Reports **352**, 219 (2001), hep-th/0001210.
[80] V. F. Kovalev and D. V. Shirkov, Journal of Physics A Mathematical General **39**, 8061 (2006).
[81] S. Saito, M. Takada, and A. Taruya, ArXiv e-prints (2009), 0907.2922.
[82] P. McDonald, H. Trac, and C. Contaldi, MNRAS **366**, 547 (2006).
[83] E. Bertschinger, Phys. Rev. D **74**, 063509 (2006), arXiv:astro-ph/0607319.
[84] J. R. Shaw and A. Lewis, Phys. Rev. D **78**, 103512 (2008), 0808.1724.
[85] S. Bharadwaj, MNRAS **327**, 577 (2001), arXiv:astro-ph/0105320.
[86] R. Scoccimarro, Phys. Rev. D **70**, 083007 (2004).
[87] B. Pandey and S. Bharadwaj, MNRAS **358**, 939 (2005), arXiv:astro-ph/0403670.
[88] Y. P. Jing and G. Börner, Astrophys. J. **617**, 782 (2004), arXiv:astro-ph/0406077.
[89] J. L. Tinker, D. H. Weinberg, and Z. Zheng, MNRAS **368**, 85 (2006), arXiv:astro-ph/0501029.
[90] J. L. Tinker, MNRAS **374**, 477 (2007), arXiv:astro-ph/0604217.
[91] J. L. Tinker, P. Norberg, D. H. Weinberg, and M. S. Warren, Astrophys. J. **659**, 877 (2007), arXiv:astro-ph/0603543.
[92] A. Slosar, U. Seljak, and A. Tasitsiomi, MNRAS **366**, 1455 (2006), arXiv:astro-ph/0507203.
[93] C. Li, Y. P. Jing, G. Kauffmann, G. Börner, S. D. M. White, and F. Z. Cheng, MNRAS **368**, 37 (2006), arXiv:astro-ph/0509874.
[94] A. Domínguez, Phys. Rev. D **62**, 103501 (2000), arXiv:astro-ph/0009414.
[95] A. Domínguez, MNRAS **334**, 435 (2002), arXiv:astro-ph/0106288.

[96] A. Domínguez, Astronomische Nachrichten **324**, 560 (2003), arXiv:astro-ph/0103313.
[97] L. M. Widrow and N. Kaiser, ApJ **416**, L71+ (1993).
[98] M. Davis and P. J. E. Peebles, ApJS **34**, 425 (1977).
[99] D. J. Eisenstein and W. Hu, Astrophys. J. **496**, 605+ (1998).
[100] C. Blake and K. Glazebrook, Astrophys. J. **594**, 665 (2003).
[101] E. V. Linder, Phys. Rev. D **68**, 083504 (2003).
[102] H.-J. Seo and D. J. Eisenstein, Astrophys. J. **598**, 720 (2003).
[103] T. Matsubara, Astrophys. J. **615**, 573 (2004).
[104] K. Glazebrook and C. Blake, Astrophys. J. **631**, 1 (2005).
[105] L. Amendola, C. Quercellini, and E. Giallongo, MNRAS **357**, 429 (2005).
[106] C. Blake and S. Bridle, MNRAS **363**, 1329 (2005).
[107] C. Blake, D. Parkinson, B. Bassett, K. Glazebrook, M. Kunz, and R. C. Nichol, MNRAS **365**, 255 (2006).
[108] V. Mukhanov, *Physical foundations of cosmology* (Physical foundations of cosmology, by V. Mukhanov. Cambridge, UK: Cambridge University Press, 2005, 2005).
[109] G. Felder and I. Tkachev, ArXiv High Energy Physics - Phenomenology e-prints (2000), hep-ph/0011159.
[110] G. Felder and L. Kofman, Phys. Rev. D **63**, 103503 (2001), arXiv:hep-ph/0011160.
[111] J. F. Dufaux, G. N. Felder, L. Kofman, M. Peloso, and D. Podolsky, Journal of Cosmology and Astro-Particle Physics **7**, 6 (2006), arXiv:hep-ph/0602144.
[112] D. Podolsky, G. N. Felder, L. Kofman, and M. Peloso, Phys. Rev. D **73**, 023501 (2006).
[113] G. N. Felder and L. Kofman, Phys. Rev. D **75**, 043518 (2007), arXiv:hep-ph/0606256.